

Relazione esercizio 2

Funzione randomLevel()

La funzione `randomLevel()` determina il numero di puntatori da includere nel nuovo nodo.

Essenzialmente la funzione `randomLevel()` restituisce 1 con probabilità 0.5, 2 con probabilità 0.25, 3 con probabilità 0.125 e così via. La distribuzione di `randomLevel()` non è quindi una distribuzione discreta uniforme, è invece una distribuzione geometrica.

La skiplist è una struttura dati probabilistica. In una skiplist implementata efficacemente i livelli più alti sono popolati in maniera meno densa rispetto a quelli più bassi. Questa caratteristica permette di svolgere ricerche più veloci. Se i livelli alti fossero popolati più densamente le prestazioni della ricerca nella skiplist si ridurrebbero. Utilizzare una distribuzione uniforme nel caso della skiplist non sarebbe efficiente proprio per questo motivo.

Studio di MAX_HEIGHT

MAX_HEIGHT	Insertion time	Correction time
3	2170.1406	0.3125
4	1047.1406	0.1718
5	367.6093	0.0469
6	132.6875	0.0313
7	61.6406	0.0156
8	30.2031	0.0000
9	13.5156	0.0000
10	6.3750	0.0000
11	3.6719	0.0000
12	2.8281	0.0000
13	2.1406	0.0000
14	2.1562	0.0000
15	1.8906	0.0000
16	1.7500	0.0000
17	1.8437	0.0000
18	1.9687	0.0000
19	1.8906	0.0000
20	1.8750	0.0000
25	1.8437	0.0000
30	1.9063	0.0000

40	1.8750	0.0000
50	1.8906	0.0000
75	1.8281	0.0000
100	1.8437	0.0000
500	1.7281	0.0000
1000	1.6562	0.0000
10000	1.5312	0.0000

Nota 1: insertion time e correction time sono entrambi in secondi.

Nota 2: l'insertion time comprende anche il tempo di leggere le parole dal file.

Con MAX_HEIGHT molto basso si nota che le prestazioni sono pessime. Questo accade perché i puntatori che permettono di “saltare” a nodi successivi della skiplist sono pochi e quindi si impiega più tempo per trovare la posizione giusta in cui inserire il nuovo elemento.

Mano a mano che aumenta il valore di MAX_HEIGHT migliorano anche le prestazioni. Aumentando MAX_HEIGHT il numero di puntatori aumenta e quando si va ad inserire un nuovo elemento è più facile trovare la posizione in cui inserirlo.

In particolare nelle prime righe della tabella si può notare che aumentare solo di uno MAX_HEIGHT comporta un dimezzamento dei tempi di inserimento.